CLAIMS

What is claimed is:

1.   An automated method for performing cross-language query translation,

comprising:

receiving a search query that includes terms in a first language;

determining possible translations of the terms of the search query into a second language;

5   locating documents in the first language that contain references that match the terms of

the search query and identify documents in the second language; and

disambiguating among the possible translations of the terms of the search query using the

identified documents to identify one of the possible translations as a likely translation of the

search query.

2.   The method of claim 1, wherein the receiving a search query includes:

presenting a graphical user interface to a user, and

receiving the terms of the search query in the first language from the user via the

graphical user interface.

3.   The method of claim 1, wherein the determining possible translations includes:

using a dictionary to identify the possible translations of the terms into the second

language.

4.   The method of claim 3, wherein the dictionary includes one or more bilingual machine-readable dictionaries.

5.   The method of claim 1, wherein the locating documents in the first language includes:

performing a search of a database of documents using the search query, and

identifying one or more of the documents in the database, each of the one or more

5   documents containing at least one reference that matches the search query and identifies a

document in the second language.

6.   The method of claim 1, wherein the references include anchor text.

7.   The method of claim 6, wherein the anchor text includes text of an anchor and text surrounding the anchor.

8.   The method of claim 1, wherein the disambiguating among the possible translations includes:

using text of the identified documents as parallel corpora, and

using a parallel corpora disambiguation technique to differentiate among the possible

5   translations of the terms of the search query.

9.   The method of claim 1, wherein the disambiguating among the possible translations includes:

determining a frequency of co-occurrence of the possible translations in the identified documents, and

5   designating one of the possible translations with a highest frequency of co-occurrence as a correct translation.

10.   A system for performing cross-language query translation, comprising:

means for obtaining a search query that includes terms in a first language;

means for performing an initial translation of the terms of the search query into a second language, the initial translation identifying one or more possible translations of the terms of the

5   search query;

means for searching a database to locate documents in the first language that contain references that match the terms of the search query and identify documents in the second language; and

means for disambiguating among the possible translations of the terms of the search

10   query using the identified documents to identify one of the possible translations as a correct translation of the search query into the second language.

11.   A system for translating search queries, comprising:

a database of documents in a plurality of languages;

a search engine configured to:

receive a search query that includes terms in a first language, and

5   search the database to locate documents in the first language that contain

references that match the terms of the search query and identify documents in a second language;

and

a query translation engine configured to:

receive the search query,

10   determine possible translations of the terms of the search query into the second

language, and

disambiguate among the possible translations of the terms of the search query

using the identified documents in the second language to identify one of the possible translations

as a likely translation of the search query.

12.   The system of claim 11, wherein the database includes a plurality of documents

distributed over a network.

13.   The system of claim 12, wherein the network is the Internet.

14.   A computer-readable medium that stores instructions executable by one or more

processing devices to perform a method for translating a search query, comprising:

instructions for obtaining a search query that includes terms in a first language;

instructions for determining possible translations of the terms of the search query into a

5   second language;

instructions for finding documents in the first language that contain references that match

the terms of the search query and identify documents in the second language; and

instructions for disambiguating among the possible translations of the terms of the search

query using the identified documents to translate the search query into the second language.

15.   A system for translating search queries, comprising:

a search engine configured to:

receive a search query that includes one or more terms in a first language, and

locate documents in the first language that contain anchor text that matches the

5   terms of the search query and identifies documents in a second language; and

a query translation engine configured to:

receive the search query,

initially translate the search query to determine possible translations of the terms

of the search query into the second language, and

10   disambiguate among the possible translations of the terms of the search query

using the identified documents in the second language to identify one of the possible translations

as a correct translation of the search query.

16.   The method of claim 15, wherein the anchor text includes text of an anchor and text surrounding the anchor.

17.   The method of claim 16, wherein the text surrounding the anchor include text of a paragraph containing the anchor.

18.   A method for performing cross-language document retrieval, comprising:

receiving a search query that includes one or more terms in a first language;

performing a search of documents in the first language to locate one or more of the first language documents that contain anchor text that matches the search query and identifies one or

5   more documents in a second language;

determining possible translations of the terms of the search query into the second language;

using the identified second language documents as parallel corpora for disambiguation among the possible translations of the terms of the search query;

10   identifying one of the possible translations as a correct translation of the search query based on the disambiguation; and

performing a search of second language documents using the correct translation of the search query.

19.   A system for performing cross-language document retrieval, comprising:

23

a search engine configured to:

receive a search query that includes one or more terms in a first language,

perform a search of documents in the first language to locate one or more of the

5   first language documents that contain anchor text that matches the search query and identifies

one or more documents in a second language;

a query translation engine configured to:

determine possible translations of the terms of the search query into the second

language,

10   use the identified second language documents as parallel corpora for

disambiguation among the possible translations of the terms of the search query,

identify one of the possible translations as a correct translation of the search query

based on the disambiguation, and

provide the correct translation of the search query to the search engine to permit

15   the search engine to perform a search of second language documents based on the correct

translation of the search query.

20.   An automated method for performing cross-language query translation,

comprising:

receiving a search query that includes terms in a first language;

determining possible translations of the terms of the search query into a second language;

5 locating documents in the first language that contain references that match the terms of

the search query and refer to other documents in the first language;

identifying documents in the second language that contain references to the other

documents; and

disambiguating among the possible translations of the terms of the search query using the

10 identified documents to identify one of the possible translations as a likely translation of the

search query.

21. A system for performing cross-language query translation, comprising:

a search engine configured to:

receive a search query that includes terms in a first language,

locate documents in the first language that contain references that match the terms

5 of the search query and refer to other documents in the first language, and

identify documents in a second language that contain references to the other

documents; and

a query translation engine configured to:

determine possible translations of the terms of the search query into the second

10 language, and

disambiguate among the possible translations of the terms of the search query

using the identified documents to identify one of the possible translations as a likely translation

of the search query.

22.   A method for performing cross-language document retrieval, comprising:

receiving a search query that includes one or more terms in a first language;

performing a search of documents in the first language to locate one or more of the first

language documents that contain anchor text that matches the search query and references one or

5   more other documents in the first language;

identifying documents in a second language that contain references to the one or more

other first language documents;

determining possible translations of the terms of the search query into the second

language;

10   using the identified second language documents as parallel corpora for disambiguation

among the possible translations of the terms of the search query;

identifying one of the possible translations as a correct translation of the search query

based on the disambiguation; and

performing a search of second language documents using the correct translation of the

15   search query.

23.   An automated method for performing cross-language query translation,

comprising:

receiving a search query that includes terms in a first language;

determining possible translations of the terms of the search query into a second language;

5   locating documents in the first language that match the terms of the search query;

identifying documents in the second language that contain references to the first language

documents; and

disambiguating among the possible translations of the terms of the search query using the

second language documents.

24.   A system for performing cross-language query translation, comprising:

a search engine configured to:

receive a search query that includes terms in a first language,

locate documents in the first language that match the terms of the search query,

5   and

identify documents in a second language that contain references to the first

language documents; and

a query translation engine configured to:

determine possible translations of the terms of the search query into the second

10   language, and

disambiguate among the possible translations of the terms of the search query

using the second language documents to identify one of the possible translations as a likely

translation of the search query.

25.   A method for performing cross-language document retrieval, comprising:

27

receiving a search query that includes one or more terms in a first language;

performing a search of documents in the first language to locate one or more of the first

language documents that match the search query;

5   identifying documents in a second language that contain references to the one or more

first language documents;

determining possible translations of the terms of the search query into the second

language;

using the identified second language documents as parallel corpora for disambiguation

10   among the possible translations of the terms of the search query;

identifying one of the possible translations as a correct translation of the search query

based on the disambiguation; and

performing a search of second language documents using the correct translation of the

search query.